IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 16, 2021, accepted December 4, 2021, date of publication December 7, 2021,
date of current version December 27, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3133809

Modeling Human Innate Immune Response Using Graph Neural Networks

SHAGUFTA HENNA, (Senior Member, IEEE)
Department of Computing, Letterkenny Institute of Technology, Letterkenny, County Donegal, F92 FC93 Ireland
e-mail: shaguftahenna@gmail.com

This work was supported by the Department of Computing and Development Office, Letterkenny Institute of Technology, Ireland.

ABSTRACT Since the rapid outbreak of Covid-19, profound research interest has emerged to understand the innate immune response to viruses to enable appropriate vaccination. This understanding can help to inhibit virus replication, prolong adaptive immune response, accelerated virus clearance, and tissue recovery, a key milestone to combat coronaviruses (CoVs), e.g., Covid-19. An innate immune system triggers inflammatory responses against CoVs upon recognition of viruses. An appropriate defense against various coronavirus strains requires a deep understanding of the innate immune response system. Current deep learning approaches focus more on Covid-19 detection and pay no attention to understand the immune response once a virus invades. In this work, we propose a graph neural network-based (GNN) model that exploits the interactions between pattern recognition receptors (PRRs)to understand the human immune response system. PRRs are germline-encoded proteins that identify molecules related to pathogens and initiate a defense mechanism against the related pathogens, thereby aiding the innate immune response system. An understanding of PRR interactions can help to recognize pathogen-associated molecular patterns (PAMPs) to predict the activation requirements of each PRR. The immune response information of each PRR is derived from combining its historical PAMPs activation coupled with the modeled effect on the same from PRRs in its neighborhood. On one hand, this work can help to understand how long Covid-19 can confer immunity for a strong immune response. On the other hand, this GNN-based understanding can also abode well for appropriate vaccine development efforts against CoVs. Our proposal has been evaluated using CoVs immune response dataset, with results showing an average IFNs activation prediction accuracy of 90%, compared to 85% using feed-forward neural networks.

INDEX TERMS Deep learning for Covid-19, modelling innate immune system, GNN for cornovirus, deep learning and immune-dependencies.

I. INTRODUCTION

CoVs patients, such as Covid-19 show unique clinical/para-clinical features including fever, cough, shortness of breath, and chest abnormalities. The features associated with chest abnormalities can be detected by medical chest imaging including computed tomography (CT) or X-ray imaging techniques [1]. These features, however, do not distinguish CoVs from pneumonia [2]. Early CoVs diagnosis is a real concern to facilitate the timely isolation of a suspected patient due to the unavailability of appropriate vaccination.

Reverse transcription-polymerase chain reaction (RT-PCR) determines the volume of specific ribonucleic acids by interacting with ribonucleic (RNA) and deoxyribonucleic acids (DNA) [3]. The test can detect severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) strain for Covid-19 detection. One of the major limitations of this test is negative at an initial stage which is detected as a positive by a CT scan [4]. Several existing works have recommended CTS scans and X-rays as a better choice due to the limited availability of RT-PCR [5]–[7].

Another popular method to screen CoVs is real-time reverse transcription-polymerase chain reaction (rRT-PCR) [8]. Although, rTR-PCR is capable to provide results within a few hours, however, its sensitivity is not high with values ranging from 37% to 71% [9]. Low sensitivity can result in a substantially high number of false-negative results at an early stage of infection. For such false-negative cases, recent works suggest chest radiology as a potential tool to detect Covid-19 [10]. Recently, the Fleischner Society has issued a consensus statement on the suitability of CT scans at an early stage in different clinical settings [11]. Owing to

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

the non-invasiveness and high sensitivity of CT scan-based Covid-19 detection, it is recommended for early Covid-19 detection and thus isolation. Radiologic diagnostic support is not available 24 hours and may have restrictions concerning location [12]. Further CT scan cannot differentiate the features of Covid-19 from the pneumonia features, thus causing uncertain predictions by radiologists.

A rapid rise in Artificial Intelligence (AI)necessitates the need for automatic Covid-19 techniques in medical imaging. Deep learning has revolutionized a rapid and early Covid-19 based on accurate analysis of chest CT images in the early stage as compared to radiologic diagnostic support. Recently, several works based on deep learning have been proposed for CoVs detection, such as Covid-19. In recent work in [13], authors proposed a deep learning model for Covid-19 called as COVID-Net. The proposed model achieved an accuracy of 83% for Covid-19 prediction. A work in [14] proposed a ChexNet model which has demonstrated outstanding results as compared to a radio-logistic diagnosis to detect pneumonia.

A reliable deep learning screening method for Covid-19 detection based on CXR images is proposed in [15]. The work consists of backbone, classification, and detection components. The classification model is based on convolutional neural networks (CNNs) with the ''sigmoid'' activation function. The anomaly module calculates anomaly scores which detects anomaly images for Covid-19 to reduce the false-positive rate. The proposal demonstrates a sensitivity of 96% with an accuracy of 71%. Until recently, CNNs have been a popular choice for pneumonia and other chest diagnoses including CoVs. Alternative simpler solutions to CNN, e.g., ResNet-50 [16] can process the images faster using numerous hidden layers.

Current efforts to CoVs detection continue to rely on deep learning-based Covid-19 detection as a potential research avenue. Most of the existing deep learning approaches are based on convolutional neural networks including ResNet, DenseNet, VCGNet which can exploit a large number of hidden layers with extensive hyperparameter tuning and long training time. Although CNN-based models are preferred deep learning models due to the processing of large data sets, however, max pooling layer can only transfer information from one layer to the next layer. It can result in the loss of small details in the data and may not capture the data which cannot be transferred from one layer to another layer. Another limitation of CNN and other deep learning models is that they cannot capture the critical relationships which exist in data. In the context of CoVs, there are dependencies between the innate human immune system and CoVs features which we can exploit to predict the CoVs detection without relying on X-ray images and CT scans.

Scarselli *et al.* [17] introduced graph neural network (GNN)approach that is widely used for predictive tasks including node classification and link prediction. GNN can learn and exploit dependencies in sparse and relational structures in data [32]. In GNN, a node can use recursive neighborhood aggregation to determine a new state and feature vector.

Unlike existing works, in this work, we model dependencies/relationships between PAMPs associated with PPRs using GNN with a particular focus on the human innate immune system. Based on this model, we propose PRRs detection and learning mechanisms that can predict PRRs activation with high sensitivity and accuracy. Our proposed PRRs prediction Algorithm PRR_{gnn} can learn useful relationships between the PRRs of the innate human immune system by using Algorithm $PRRs_{learn}$. Experimental evaluations and results based on PRRDB2.0 considering a positive dataset and Swiss-Prot as negative data set show that the proposed Algorithm PRR_{gnn} can predict PRRs activation with high sensitivity and accuracy compared with the PRRs activation traditional feed-forward neural network (FNN). Our evaluations reveal that the PRR_{gnn} considers useful dependencies between PRRs for IFNs activation. Our results reason that the proposed PRR_{gnn} can extract useful insights into PRRs interactions that can be useful for CoVs detection with high accuracy at an early stage.

The remainder of this paper is organized as follows. Section II investigates the existing works relevant to PRRs activation using deep learning. Section III details the immune response features. Section IV presents and discusses the proposed CoVs immune response system model based on GNN and Algorithms with learning. Section V presents the results and discussion of the proposed method. Section VIdiscusses the limitations of the study and conclusions and future works are given in Section VII.

II. RELATED WORKS

The last few months have witnessed several Covid-19 investigations relevant to its detection and spread [18]. The Covid-19 detection has been addressed based on different deep learning approaches using medical imaging including CT scans and chest X-ray. Other works discuss the spread of Covid-19 with a focus on the number of confirmed cases, recoveries, and deaths. Recent years have seen a rise in deep learning-based solutions to disease diagnosis, primarily based on x-ray images. In a similar effort, a deep neural network model called ChexNet for pneumonia detection based on chest x-ray images is proposed in [19]. The model demonstrated exceptional performance results in terms of accuracy. Following this work, another model called ChestNet [20], a deep learning solution to predict thorax disease based on chest x-ray images.

In a recent work on Covid-19 diagnosis, authors have evaluated various convolutional neural networks (CNN) coupled with a pre-trained ResNet 50 model with 98% accuracy [5]. The evaluation study classifies healthy and Covid-19 infected patients. The work reports Covid-19 diagnosis with an accuracy of 97% using InceptionV3 and 87% based on Inception-ResNetV2. The study, however, does not consider the discrimination between pneumonia conditions from Covid-19. In another work [21], authors

TABLE 1. Comparison of machine learning techniques used for protein sequence analysis.

Model	Advantages	Limitations
SVM [26]	• Efficiently handles non-linear data. • Demonstrates good generalization capabilities. • Stable under new data points.	• Slower training speed. • Not suitable for larger datasets, e.g., protein interactions. • High computational complexity.
RF [26]	• Efficiently works with colinearity. • Works efficiently with high number of features in the data. • Highly stable under new data points.	• Low generalization capabilities. • High complexity for larger protein datasets. • Training time increases as the number of trees increases.
FNN [27]	• Works well with both linear and non-linear data. • Fault tolerant under high regularization. • Scales well to larger datasets. • Good generalization ability.	• Unable to leverage underlying structure in PRRDB. • Not powerful to discover new patterns in molecular networks. • Unable to model PRRs relationships efficiently.
KNN [28]	• Suitable for non-linear decision boundaries. • Works well with high number of features in the data. • Highly stable under new data points. • Little to no explicit training requirements.	• Slower prediction speed for large datasets. • Computationally expensive with high memory requirements. • Highly sensitive to outliers in PRRDB dataset.

presented a CNN-based solution to Covid-19 diagnosis from chest x-ray images. The proposal suggested a classification accuracy of approximately 97% with MobileNet.

Wang *et al.* [13] proposed COVID-Net, a deep learning model for Covid-19 detection. The model demonstrated an accuracy of 83.5% for Covid-19 classification based on healthy, bacterial-infection, and viral-infection classes. In another work in [22], authors evaluated different deep learning models for Covid-19 detection using chest x-ray images. The authors also proposed a COVIDX-Net model based on 7 CNN models. In a similar work on Covid-19 detection [23], authors have classified and evaluated the chest x-ray images as healthy or infected using various deep learning approaches, i.e., AlexNet, VGG16, GoogleNet, ResNet-101, and Inception-ResNet-v2, etc. The evaluations demonstrated an accuracy of approximately 95% for RestNet50.

In [24], Hassanien *et al.* proposed a multi-level threshold-based support vector machine (SVM) framework for classification. The proposed system detects Covid-19 using x-ray images. It relies on 40 x-ray images of 15 healthy and 25 infect cases. It has demonstrated an accuracy of 97% with a sensitivity of 96%. In a similar work in [25], the authors demonstrated deep learning algorithms for Covid-19 detection based on CT scans of 157 patients. The system evaluated the deep learning-based detection algorithm using two subsystems. One subsystem considers a 3D analysis, whereas the second one performs a 2D analysis of the CT scans using a Resnet-50-2 with an area under the curve of 99%. The subsystem demonstrated a sensitivity of 98% with a specificity of 92%.

A work in [26] evaluated SVM and random forests (RF) for predicting interleukin-10 (IL-10) inducing peptides, a cytokine responsible for suppressing the immune system. Both the models were trained on 394 IL-10 inducing and 848 non-inducing peptides. Results in the study reveal that the RF performs better in terms of accuracy in contrast to SVM using dipeptide composition.

A study in [27] evaluates an artificial neural network/feed-forward neural network to predict hemolytic activity from a peptide's primary sequences. Results show that the classifier demonstrates a validation accuracy of 85.7% and outperforms SVM and RF. In another work, [28], the authors used various machine learning models including k-nearest neighbour (KNN), linear SVM, and RF for the prediction of proteomes as new B-cell epitopes in vaccine design using reference epitope sequences. The models are evaluated under peptide sequences that are converted into molecular descriptors. The non-linear SVM uses non-linear interactions between the molecular properties for proteomes predictions.

In recent work, Wang *et al.* [29] proposed a deep learning solution to predict Covid-19 based on 195 regions of interest (ROIs) of 395 $\times$ 223 to 636 $\times$ 533 pixels from the CT scans of 44 Covid-19 positive patients and 258 ROIs from 50 Covid-19 negative patients. The internal validation of the proposed model has suggested an accuracy of 83% with the specificity of 80.5% and sensitivity of 84%. The proposed inception model demonstrates an accuracy of 73.1% with a sensitivity of 74% and a specificity of 67%. In [30], authors have proposed deep ResNet-50-based classification system for lung diseases including Covid-19 and pneumonia. The model is trained on more than 60K CT scans from

918 patients including Covid-19 and non-Covid-19. The proposed model has demonstrated an accuracy, specificity, and sensitivity of approximately 98%. Another similar work in [31] reports image classification as Covid-19 and non-Covid-19 using RestNet-based methods. However, contrary to the work in [30], this work uses a Bayesian function to differentiate Covid-19 and non-Covid-19 images with an accuracy of 86.7%.

Table 1 presents a comparison of recent studies that use machine learning models to predict IL-10 inducing peptides responsible for immune response, hemolytic activity from peptide sequences, and proteomes as new B-cell epitopes in vaccine design. These approaches, however, do not model the interactions among PRRs as a graph to initiate a cascaded activation sequence for immune response. This study is the first effort to model interactions among PRRs as a graph to understand the human innate immune response and make predictions using graph-based deep learning.

III. IMMUNE RESPONSE FEATURES:PATHOGEN-ASSOCIATED MOLECULAR PATTERNS (PAMPs)

The primary features with respect to human immune system are PAMPs that are recognized by different pattern recognition receptors (PRRs). Various PAMPs features associated with different PRRs are discussed below.

A. TOLL-LIKE RECEPTORS PAMPS

Toll-like receptors (TLRs) can recognize PAMPs denoted as $PAMPs_{tlrs}$ including lipids, proteins, lipoproteins, and nucleic acids. The recognition of $PAMPs_{tlrs}$ occurs in cell membranes, endosomes, lysosomes, endocytolysosomes, and other locations in cells [33]. Different TLRs can induce different biological responses via subsequent activation of varied adapter proteins, such as MyD88, TIRAP, TRIP, and TRAM. All these adapter proteins share the Toll receptor (TIR) structure. MyD88 is the first identified TIR family member which acts as an adapter protein by almost all TLRs except TLR3 [34].

B. RIG-I-LIKE RECEPTORS PAMPs

RIG-I-like receptors (RLRs) can recognize PAMPs features based on nucleic acids denoted as $PAMPs_{rig}$. $PAMPs_{rig}$ can result from different infections including Influenza A virus (IAV), Measles virus (MV), and Hepatitis C virus (HCV). These $PAMPs_{rig}$ can also include viral nucleocapsid proteins containing triphosphate and double basic acid RNA at the 5-end [35]. These features are used to identify RNAs of picornaviruses, including poliovirus (PV) and Encephalomyocarditis virus (EMCV). These RNAs are primarily characterized by long double-stranded RNA of more than 1 kbp.

C. NUCLEOTIDE-BINDING AND OLIGOMERIZATION DOMAIN-LIKE RECEPTORS ((NLRs) PAMPS

NLRs recognize PAMPs features $PAMPs_{nlrs}$ based on conserved Nucleotide-binding and oligomerization domain (NOD) structure. These features are based on various proteins/complexes called the inflammasome, reproduction, and regulatory NLRs. The inflammasome consists of at least eight NLR proteins, including NLRP1, NLRP3, NLRP6, NLRC4, NLRC5W, and AY2 [36], [37].

D. C-TYPE LECTIN-LIKE RECEPTOR (CLRs) PAMPs

CLRs recognize PAMPs features, denoted as $PAMPs_{clrs}$. CLRs are activated directly through macrophage-induced Mincle and CLEC4E, and Dectin-2 CLEC6A) receptors. The indirect activation of CLRs is triggered by the HAM-like motifs in the intracellular tail of the receptor, e.g., Dectin-1 (CLEC7A) and DNGR-1 (CLEC9A) [37]. These activations result in acidified spleen tyrosine kinases which trigger the formation of CARD9, B-cell lymphoid tissue 10 (BcL10), and Maltlcomplex formation. The signaling pathways can also include SyK and JNK acidified apoptosis-related protein granule, e.g., ASC [38]. These pathways, in turn, activate downstream *PAMPs*, i.e., NF-kB and MAPKs. These *PAMPs* can trigger various cellular responses, e.g., phagocytosis, maturation, and chemotaxis of cells [39].

E. CYTOPLASMIC DNA RECEPTOR PAMPs

CLRs Cytoplasmic DNA receptor (CDR) can recognizes DNA CpG islands, denoted as $PAMPs_{cdr}$ [40]. Examples of CDRs include AIM2-like receptors (ALRs), DNA-dependent activator of IFN-regulatory factor (DAI), leucine-rich repeat flightless-interacting protein 1 (LRRFIP1), DExD/H-box RNA helicase (DDX), Meiotic recombinant protein 11 Homolog A (MRE11), RNA polymerase III (Pol III), DNA dependent protein kinase (DNA-PK), DNA repair-related proteins Rad50 and Sry-related HMG box 2 (Sox2) [41]. DAI recognizes PAMPs based on Z-type DNA and B-type DNA [43]. These DNAs depend on the length of DNA. AIM2-related PAMPs recognize double-stranded DNA. IFI16 and cGAS receptors can recognize cytosolic DNA recognition and are capable of type I interferon [42].

F. TYPE I INTERFERONS

Pathogen-associated molecular patterns recognize the viral nucleic acid, activate IRF3 and IRF7, and promote type I interferons (IFNs). IFNs trigger the JAK-STAT signal pathway, thereby promoting IFN-stimulated genes (ISGs) [44], [45]. IFNs are antiviral molecules that contribute a major role to immunomodulatory. Specifically, antigens resultant from these IFNS restricts infected target and T/B cells and any blockage to IFNs can affect the survival of the virus [46], [47].

PRRs consist of three types including membrane, secretory and cytoplasmic [48]. The membrane PRRs include TLR2, TLR4, mannose receptor (MR), and scavenger receptor (SR). On the other hand, the secretory PRRs consist of mannose-binding lectin (MBL) and C-reactive protein (CRP). TLR3, TLR7/8, and NLRs form cytoplasmic PRRs. Among all these types, PRRs including TLRs, RLRs, and NLRs result in IFN production. SARS-CoV and other coronaviruses are pathogenic and are sensitive to IFN-a/b. The N-protein of

SARS-CoV is classified as an immune-escape-protein and is an antagonist against host interferon response [49].

A well-reported EV71 infection down-regulates JAK1, p-JAK1 and p-TYK2. This down-regulation results in blockage of JAK-STAT signaling pathway thereby inhibits IFNs production. Reduced function of IFNs results in high EV71 replication in host cells [50]. Similarly, Ebola virus (EBOV) inhibits IFNs production by promoting cytokine signal inhibitory factor-1 (SOCS1) which also blocks JAK-STAT signaling pathway [51]. Further, influenza A is also capable to inhibit IFN-I production by activating SOCS3 [52].

G. DENDRITIC CELLS

Dendritic cells (DCs) stimulate the activation of T-lymphocytes and B-lymphocytes and play a vital role in innate and adaptive immunity. Mature DCs can activate T cells, thereby directly affecting the adaptive immune responses. DC precursor cells are differentiated based on inducers including GM-CSF, IL-4, and TNF-α if not transfected with HIV-1 Nef protein. The viral antigens capability of DCs is limited by the HIV-1 which reduces the major histocompatibility antigen I (MHC I) on DC's surface. The functionality of DCs interferes with the viral infections which helps the viruses to evade the adaptive immune response of the host [53].

H. DEFENSINS

Defensins consist of antibiotic peptide molecules which are critical for the host's innate defense system. It eliminates bacteria including viruses, fungi, and tumor cells. Defensins are found in neutrophils that consist of small molecular cationic polypeptides. Defensin HNP-1 inactivates viruses including HSV-1, HSV-2, cytomegalovirus (CMV), VSV, and IAV [54]. There are some studies that reveal that the human neutrophil defensin (HNP1-3) is not able to inhibit or kill SARS-CoV [55].

IV. METHODS

A. GRAPH NEURAL NETWORK MODEL

We formulate the PRR activation as a neighbor-dependency problem. The proposed solution called PRR_{gnn} uses GNN [17] approach to model the dependencies between the PRRs. These PRRs dependencies can contribute towards the innate immune response prediction. Given a human innate immune system with various PRRs that recognize different PAMPs, the use of GNN is motivated by the fact that an immunity-response against CoVs results in inter-PRRs dependencies explaining the impact of CoVs on the health of a particular individual.

A trained PRR_{gnn} can predict the role of innate immune response based on the PAMPs features of a PRR, and the PAMPs features of neighboring PRRs.PRRs are assumed to have PAMPs activation dependencies in the host innate immune system. PRR_{gnn} consists of two feed-forward neural networks (FNNs). PRR_{gnn} uses the first FNN to compute its next activation state denoted as a_n based on the PAMPs features of its neighboring PRRs n^*. The second FNN predicts the activation of PRR based on the a_n and its historical PAMP features. PRR_{gnn} model for PRR activation prediction is given below in form of Equation 1 and Equation 2.

$$a_n = \sum_{i \in n} h_w(f_n, f_i, a_i), \quad \forall n \tag{1}$$

$$C_n(t) = g_w(a_n, f_n), \quad \forall n \tag{2}$$

The equation 1 represents that the activation state function of a PRR depends on the PAMPs features f_i and its activation state, i.e., a_i. In the equation, h_w represents a parametric function that depends on its PAMPs features, PAMPs features of its neighboring PRRs, and the activation state of neighboring PRRs. The g_w function in Equation 2 denotes the activation prediction of PRR based on the activation state a_n from equation 1 and its historical PAMPs. PRR_{gnn} consists of both h_w and g_w function for each of the PRRs that converge exponentially fast. The fast convergence process yields stable PRR and predicted PRR activation states to predict the innate immune response.

B. PRRs GRAPH NEURAL NETWORK MODEL (PRR_{gnn})

Given a host innate immune system which can activate PRRs to identify PAMPs to detect viral infections, the PRR_{gnn} aims PRRs activation to predict immune response. The immune response of a PRR impacts the activation of neighboring PRRs that reflects critical relationships among PRRs. It is possible to model these immune response PRRs dependencies with a GNN which can assist a faster and accurate innate immune response prediction that can be used to treat the patient promptly and helps to prioritize critical cases. These inter-dependencies among PRRs reflect the activation dependencies of the immune response system.

GNN-based immune response model, i.e., PRR_{gnn} is explained with help of Equation 1 and Equation2. Figure 1 shows an example of activation-related dependencies among PRRs. Figure 2 shows different modules of PRR_{gnn}. It consists of immune features, immune states, one FNN for h_w, and the other for g_w. The h_w function provides the PAMPs features of a PRR and its state, whereas, the g_w function predicts the activation state of PRR, e.g., IFNs based on its past features and state combined with the neighbouring PRR 's PAMPs features.

C. PRR_{gnn} FEATURES

PRR_{gnn} features include PAMPs recognized by different PRRs. These PRRs include TLR, RLR, NLR, CLmin, cGAS, IFI16, STING, DAI, etc. Few of these have been discussed in detail in Section III. These PAMPs features are provided as an input to predict the next activation state of PRR a_n by using Equation 1. The output function in Equation 2 takes the PAMPs of the current PRR f_n and the a_n. For example, PAMPs recognized by TLRs $PAMPs_{tlrs}$ include lipids, lipoproteins, proteins, and nucleic acids of the bacteria and viruses.

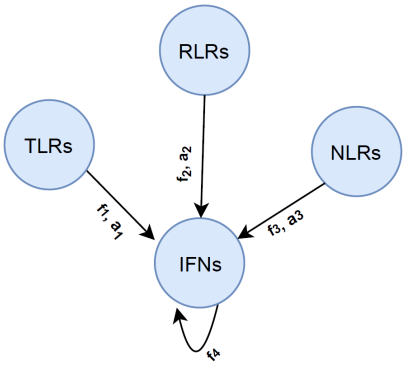

FIGURE 1. States & Features from IFNs's neighbouring PRRs.

D. PRR_{gnn} STATES

The PRRs are assumed to preserve the activation state a_n derived by the parametric function h_w. This function takes the PAMPs features of a PRR and PAMPs of neighboring PRRs. h_w shows that the activation state of a PRR depends on the PAMP features and PAMP features of its neighboring PRRs, thereby reflecting the activation dependencies of the immune response system.

The function h_w in Equation 1 computes the next activation state a_{i+1} of PRR based on the current state a_i. It means, the current state $a_n(i)$ of PRR depends on the previous states $a_m(i-1)$ of its neighbour m as given in Equation 5.

$$a_n(i+1) = \sum_{m \in n*} h_w(f_n, f_m, a_m(i)) \tag{3}$$

Figure 1 depicts the activation dependency of a PRR on the features of neighbouring PRRs. The figure shows that the state of the PRR IFNs depends on the activation states a_1, a_2, a_3 and features of TLRs, NLRS, and RLRs, i.e., f_1, f_2, f_3. The activation state of IFNs denoted as a_4 can be predicted based on Equation 1. The state a_4 at IFNs in figure 1 is given as Equation 4.

$$a_4 = h_w(f_4, f_1, a_1) + h_w(f_4, f_2, a_2) + h_w(f_4, f_3, a_3) \tag{4}$$

E. OUTPUT FUNCTION

PRR activation based on the PAMPs features depends on the next activation state as computed by the h_w function and a PRR ś features. For the example in figure 1, the output function is based on a g_w as given in Equation 5. The function g_w is an FNN trained by the gradient descent method discussed later. The g_w function of each PRR predicts its next activation state.

$$C_4(t) = g_w(a_4, f_4) \tag{5}$$

F. PRR_{gnn} ALGORITHM

The PRR_{gnn} is illustrated with the help of Equation 1 and Equation 2. It takes PAMPs features of a PRR as input and outputs its activation state. The model computes the activation state of each PRR. This computation is the result of an iterative process where each PRR has h_w functions equivalent to the number of its neighboring PRRs in the innate immune system. The PRR activation contributes to determining the innate immune response at an early stage.

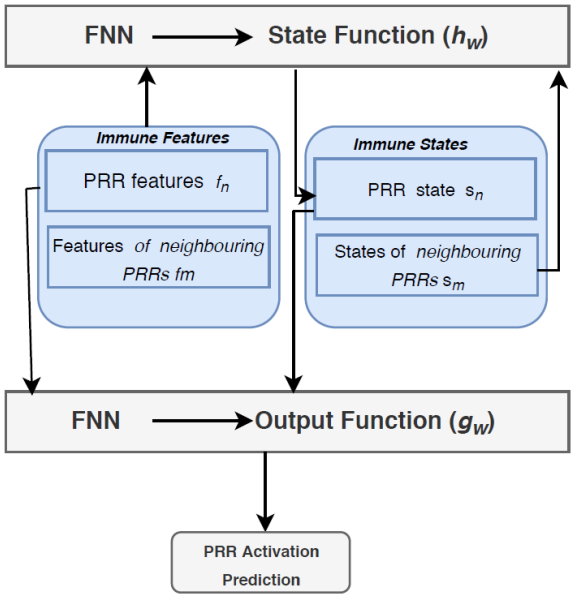

FIGURE 2. PRR activation using PRR_{gnn}.

Both the h_w and g_w functions are implemented as FNN. This is illustrated in Algorithm 1. The algorithm shows that the PRR activation prediction is based on:

1) Observation of a PRR's PAMPs and its directly connected neighboring PRRs,
2) Computation of the next activation state of PRR based on the observed PAMPs features, and,
3) Use PAMPs observations from step 1 and activation state from step 2 to predict the IFNs activation that can help to determine the innate immune response at an early stage.

G. PRR_{gnn} LEARNING

Both the FNNs for the two functions h_w and g_w require training in order to make accurate predictions. This is accomplished with the help of input Q and output P which is used to tune the weights of the FNNs. In our use-case of Covid-19, this requires a dataset for a human innate immune system including historic and current PAMPs features. The learning of FNNs minimizes the cost function as given in Equation 6.

$$q_x = \sum_{m \in M} \left(\frac{1}{2}(C_n - P_n)^2 + \beta L(C_n) \right) \tag{6}$$

Equation 6 computes the error to minimize the cost function. The loss is added to the error function and can be scaled based on β. The value of L penalizes the weights of the FNN whenever its output, i.e., C_n exceeds the μ [17]. Algorithm 2 illustrates the learning algorithm based on gradient descent. The algorithm adjusts the weights for both

Algorithm 1 PRR_{gnn} Algorithm

Input Data: $k = 0$, state(k)=0,number of iterations M
Ouput: Predicted PRRs Activation ▷ To determine immune response

1: $PAMPs_features \leftarrow PAMPs_features(PRRs) \cup PAMPs_features(neighbouring(PRRs)$ ▷ Observe PAMPs features for all PRRs and their neighbouring PRRs
2: **while** $k < M$ **do**
3: $\quad a_n(k+1) = \sum_{m \in n} h_w(f_n, f_m, s_m(k)), \forall n$
4: $\quad k \Leftarrow k+1$
5: $C_k = g_w(s_n(k), f_n), \forall n$
6: $PRRs_a \Leftarrow C_k$ ▷ Predict PRRs activation
7: return (PRR_a, a_n)

Algorithm 2 $PRRs_{learn}$ Algorithm

Input Data: $w = 0$, k, Desired criteria ACC
Ouput: Learned FNNs for h_w and g_w

1: **while** $k < ACC$ **do**
2: $\quad a_n(k) \leftarrow$ Algorithm 1 (k)
3: $\quad C_k \leftarrow$ Algorithm 1 (k)
4: $\quad \frac{\partial q_x}{\partial w} \leftarrow$ BPPgradient(q_x, w)
5: $\quad w(k+1) \leftarrow w(k) - \alpha \frac{\partial q_x}{\partial w}$
6: $\quad k \Leftarrow k+1$
7: return learned h_w and g_w

the FNNs for each h_w and g_w in order to minimize the cost functionq_w given in Equation 6.

The Algorithm 2, $PRRs_{learn}$, tunes the weights w of FNNs based on gradient descent, each for h_w and g_w to minimize the cost given in Equation 6. The algorithm executes iteratively and updates the state and output as given in line 2 and line 3 by calling the Algorithm 1. Line 4 computes the gradient $\frac{\partial q_x}{\partial w}$ of the cost function for h_w and g_w for weight w. Line 5 updates the weight as $w(k+1) \leftarrow w(k) - \alpha \frac{\partial q_x}{\partial w}$ where α denotes the learning rate. Line 4 uses the backpropagation-through-time (BPTT) [17] to compute the gradient of the cost function for both the FNNs each for h_w and g_w.

V. RESULTS AND DISCUSSIONS

A. EXPERIMENTAL SETUP AND DATASET

PRRs sequences are obtained from the database PRRDB2.0 [56] that constitute a positive dataset. PRRDB2.0 provides extensive information about unique PRRs and PAMPs. Each entry of the PRR in PRRDB2.0 has details of the PAMPs features. The main aim of the dataset is to understand human innate immunity. We have considered two datasets, i.e., a positive and a negative. For the positive dataset, initially, approximately 2727 PRRs are selected randomly. From these PRRs, only 179 PRRS are considered after eliminating redundant sequences. The negative dataset is derived by selecting sequences randomly from the Swiss-Prot [60], constituting non-pattern recognition receptors(Non-PRRs). To formulate positive and negative clusters, we have considered a cutoff of 40% sequence similarity without significantly reducing the number of proteins from the dataset for unbiased models training and testing of FNN and GNN. [57], [58] Based on the selected cutoff, we have obtained 106 positive and 210 negative subsets to train and test the PRR_{gnn} model.

Experiments are based on PyTorchLightning to implement the GNN and FNN models using the 'LightningModule' [61].The proteins are converted to fixed vectors to facilitate training and testing of the models, and the model hyper-parameters are tuned by using the parameter function of the implemented 'LightningModule', i.e., GNN and FNN. The training considers 4 positive and 4 negative subsets from 106 positive and 210 negative subsets. The test set, however, considers one positive and one negative subset to evaluate the model performance. From the 'LightningModule', fit() function is used to train the models, and forward() for inference/predictions.

Experiments consider a 5-fold cross-validation technique to address the issues of bias for the positive and negative classes. For each fold, the proteins in the test set are matched with the training set using a cutoff of 40%sequence similarity. This process is repeated 5 times to cover all the proteins from the training set. Training and test sets are based on positive and negative subsets. 4 positive and 4 negative subsets are combined for the training set, whereas, one positive and one negative subset forms the test set.

This process is repeated 5 times. In each fold, the positive and the corresponding negative subset acts as a test set exactly once. The experiments consider the 1000 number of trials to understand the learning curve and convergence of both the GNN and FNN in terms of accuracy, sensitivity, specificity, and AUC. These trials also signify that the model is not overfitted.

B. PERFORMANCE METRICS

PRR_{gnn} evaluation is based on threshold-dependent and threshold-independent scoring parameters. Threshold-based scoring is used to calculate sensitivity, specificity, accuracy, and Matthew's correlation coefficient (MCC). Sensitivity is defined as correctly predicted positive PRRs divided by the total positive PRRs. On the other hand, specificity denotes the true negative rate of PRRs. Accuracy represents a model's ability to predict true positive PRRs. MCC calculates the correlation coefficient between actual and predicted PRRs. The threshold-independent performance metric area under the receiver operating characteristic curve (AUC) represents the plot between sensitivity and false-positive rate.

C. ANALYSIS OF RESULTS

In the experiments, the trained PRR_{gnn} system is tested to determine its sensitivity, specificity, prediction accuracy, MCC, and AUC over $1,000$ tests. The evaluation results for the PRR activation based on FNN are given in Figure 3 to Figure 7. Figure 3 shows the sensitivity of the FNN-based PRR activation. The trained traditional neural network achieves 78% sensitivity and specificity of

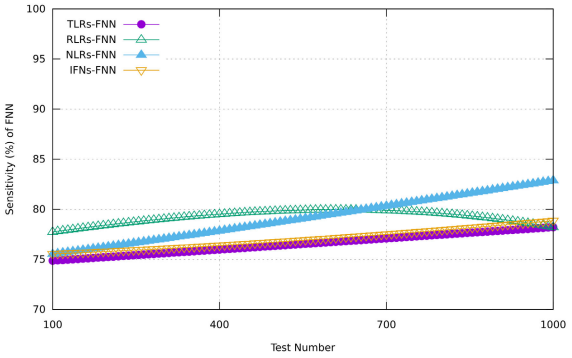

FIGURE 3. Sensitivity of FNN.

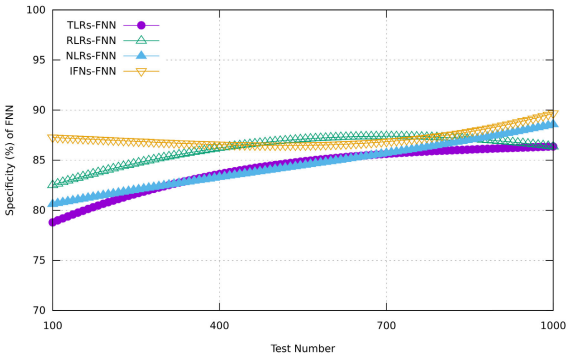

FIGURE 4. Specificity of FNN.

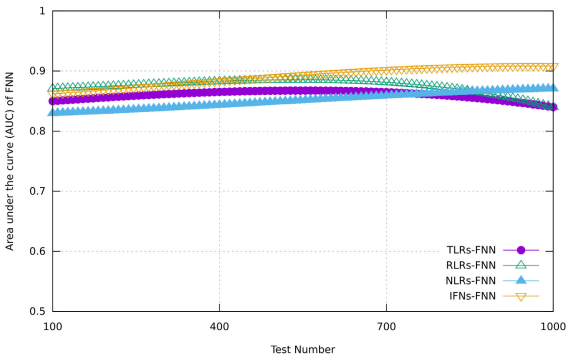

FIGURE 5. AUC of FNN.

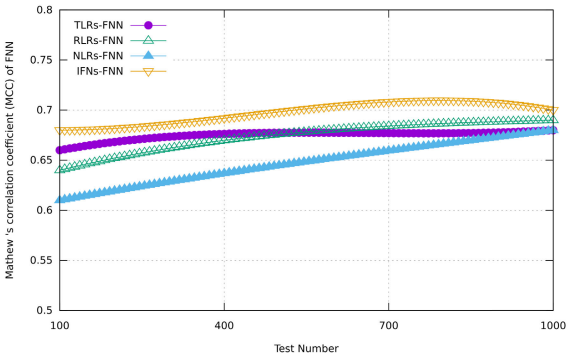

FIGURE 6. MCC of FNN.

approximately 89% for the activation of IFNs to predict innate immune response against CoVs. The sensitivity for

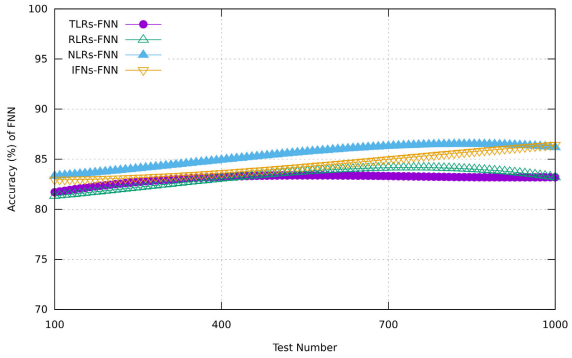

FIGURE 7. Accuracy of FNN.

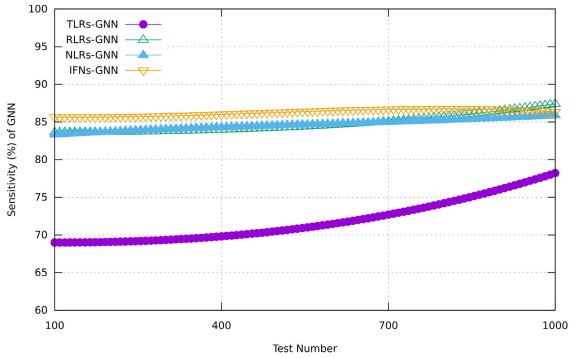

FIGURE 8. Sensitivity of PRR_{gnn}.

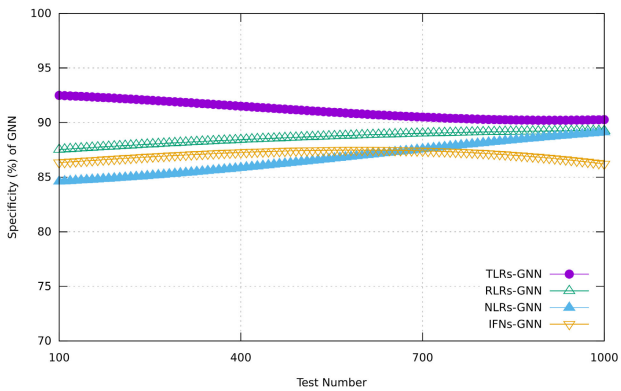

FIGURE 9. Specificity of PRR_{gnn}.

the NLRs activation is 83% with a specificity of 89%. TLR's sensitivity associated with the TLRs using FNN varies 74% to 76% for all the test measurements with a specificity of a maximum 83%. PRR_{gnn}-based IFNs activation demonstrates an approximately 86% sensitivity and 87% specificity. The sensitivity of the TLRs activation based on FNN varies between 75% to 77%, with specificity between 76% to 83%. The PRR_{gnn} shows NLRs activation sensitivity with a value 84% for all the tests, with maximum specificity of 88%. The maximum accuracy achieved by the FNN-based IFNs activation is approximately 85%, with AUC 0.9%, and MCC 0.74%. On the other hand, PRR_{gnn}-based IFNs activation accuracy is 90% for all the tests, with AUC 0.9%, and MCC 0.73%.

TABLE 2. Performance comparison of GNN with other machine learning classifiers techniques.

Model	Sensitivity	Specificity	Accuracy	AUC	MCC
PRR_{gnn}	86.35	86.21	90.41	0.74	0.89
SVM	77.02	80.90	79.62	0.731	0.58
FNN	78.85	89.67	86.44	0.707	0.908
RF	71.23	77.10	74.95	0.682	0.58
KNN	69.21	73.23	70.59	0.673	0.52

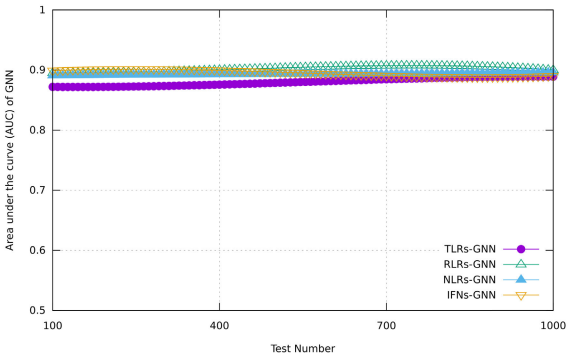

FIGURE 10. AUC of PRR_{gnn}.

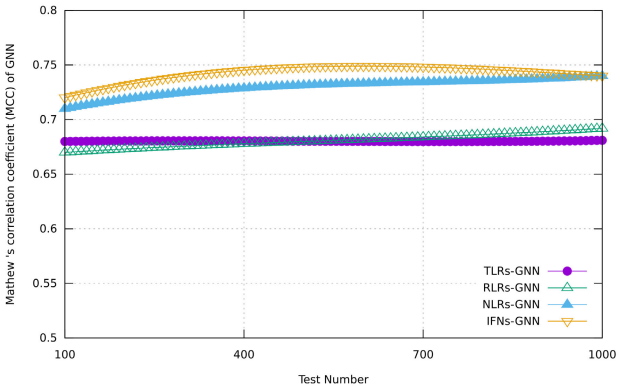

FIGURE 11. MCC of PRR_{gnn}.

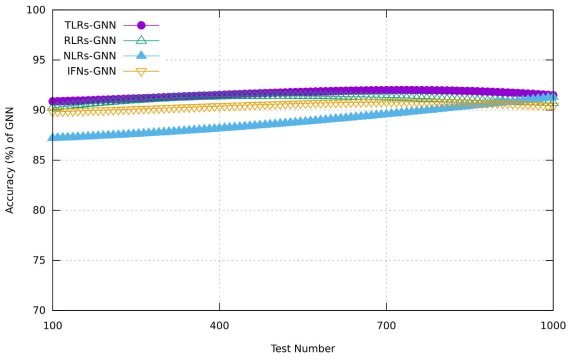

FIGURE 12. Accuracy of PRR_{gnn}.

Since the GNN model can specifically capture the features of the neighbouring PRRs to predict PRR activation, it performs better than FNN in terms of accuracy, specificity, and sensitivity. Given that a PRR 's activation can be predicted as a combined effect of PAMPs of its neighboring PRR and its own PAMP, the activation prediction accuracy and sensitivity is better than the FNN. We can observe from

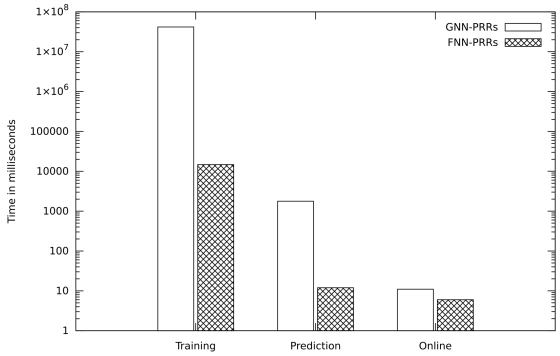

FIGURE 13. Training and prediction time of PRR_{gnn}.

the Figure 3, Figure 8, Figure 4, and Figure 9 that for the IFNs, sensitivity, and specificity are nearly equal which is an ideal result for the PRR_{gnn} PRR activation model. The AUC achieved by FNN-based activation for RLR, NLR, and TLR is approximately 0.85, while the PRR_{gnn} achieves an AUC of 0.9 for these PRRs as shown in Figure 5 and Figure 10. This indicates the significance of our PRR_{gnn} Algorithm for the action of PRR to predict the innate immune response against Covid-19. It can be observed from Figure 6 and Figure 11 that the MCC of FNN-based IFNs activation is 0.7. As compared to FNN-based IFNs activation, the GNN model shows higher performance with an MCC of nearly 0.75. The PRR_{gnn} NRLs activation demonstrates an MCC of 74% as compared to the FNN-based method with an MCC value of 0.74. Figure 12 shows that the accuracy achieved by the PRR_{gnn} is significantly better than the FNN-based PRR activation, i.e., above 90% for all the test cases.

Figure 13 compares the training time of PRR_{gnn} PRRs activation with the FNN-based method. The figure also shows the time taken by both neural networks to make predictions. For both cases, GNN takes more time as compared to the FNNs. This additional time is attributed to the fact that GNN consists of multiple FNNs. It is evident from Figure 13 that the GNN training time takes about 41,000 seconds. The GNN training process is offline that does not affect the online performance of the system to make predictions. Once trained, the weights of the FNNs are saved and loaded to make predictions. There is not a significant time difference for online predictions, e.g., for the time-critical system's predictions are kept running. These online predictions take about 11ms that is comparatively less than the offline training and prediction time.

From the above results, it is evident that the PRR_{gnn}-based PRRs activation outperforms the FNN-based activation

method. This benefit is due to the neighborhood awareness of the GNN method. Another characteristic of GNN is its multiple FNNs with a higher number of layers, each with a higher number of neurons.

Table 2 presents the comparison of with FNN and other machine learning classifiers, such as SVM, RF, and KNN using PRRDB 2.0 dataset for 1000 trials to predict IFNs activation.

VI. LIMITATIONS OF THE STUDY

The PRRDB2.0 and Swiss-Prot datasets are not large enough to train a reliable version of the PRRgnn model. There is a margin to improve the size and the quality of these datasets through public review to avoid bias. Further, it will be interesting to go beyond neighboring PRR aggregation to come with powerful architecture to enable PRRgnn learning for human innate immune response with a focus on generalization properties of GNN [59].

VII. CONCLUSION AND FUTURE WORK

This work proposes a graph neural network-based approach to predict the activation of PRRs in a human innate system to evaluate the immune response against CoVs. The proposed model exploits the interactions between PRRs inherent to a human innate immune response system. In comparison to FNN-based IFNs activation, the proposed model relies on dependencies between different PRRs to predict activation of IFNs. Our work can act as a milestone to investigate graph-based analytic to predict infectious disease through an understanding of the human innate immune response system. Future work aims to consider edge features of a graph in a GNN to weigh the importance of each PRR interaction along with the PAMPs features. Further, these interactions can be modeled with self-supervised and semi-supervised learning to predict PRRs activations of innate human immune response.

VIII. DATA AVAILABILITY

The datasets used for this study can be found at the PRRpred webserver (https://webs.iiitd.edu.in/raghava/prrpred/ dataset.php).

REFERENCES

[1] C. Huang, Y. Wang, X. Li, L. Ren, and X. Gu, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *Lancet*, vol. 395, pp. 497–506, Feb. 2020.

[2] M.-Y. Ng, E. Y. P. Lee, J. Yang, F. Yang, X. Li, H. Wang, M. M.-S. Lui, C. S.-Y. Lo, B. Leung, P.-L. Khong, C. K.-M. Hui, K.-Y. Yuen, and M. D. Kuo, "Imaging profile of the COVID-19 infection: Radiologic findings and literature review," *Radiol., Cardiothoracic Imag.*, vol. 2, no. 1, Feb. 2020, Art. no. e200034.

[3] *Laboratory Testing for Coronavirus Disease 2019 (COVID-19) in Suspected Human Cases: Interim Guidance, 2 March 2020*, World Health Org., Geneva, Switzerland, 2020.

[4] T. Ai, Z. Yang, H. Hou, C. Zhan, C. Chen, W. Lv, Q. Tao, Z. Sun, and L. Xia, "Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases," *Radiology*, vol. 296, no. 2, pp. E32–E40, Aug. 2020, doi: 10.1148/radiol.2020200642.

[5] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks," 2020, *arXiv:2003.10849*.

[6] H. S. Maghdid, A. T. Asaad, K. Z. Ghafoor, A. S. Sadiq, and M. K. Khan, "Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms," 2020, *arXiv:2004.00038*.

[7] S. U. K. Bukhari, S. S. K. Bukhari, A. Syed, and S. S. H. Shah, "The diagnostic evaluation of convolutional neural network (CNN) for the assessment of chest X-ray of patients infected with COVID-19," *medRxiv*, 2020.

[8] M. J. Loeffelholz and Y.-W. Tang, "Laboratory diagnosis of emerging human coronavirus infections—The state of the art," *Emerg. Microbes Infections*, vol. 9, no. 1, pp. 747–756, Jan. 2020, doi: 10.1080/22221751.2020.1745095.

[9] T. Ai *et al.*, "Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases," *Radiology*, vol. 296, no. 2, Feb. 2020, Art. no. 200642, doi: 10.1148/radiol.2020200642.

[10] D. Li *et al.*, "False-negative results of real-time reverse-transcriptase polymerase chain reaction for severe acute respiratory syndrome coronavirus 2: Role of deep-learning-based CT diagnosis and insights from two cases," *Korean J. Radiol.*, vol. 21, no. 4, pp. 505–508, 2020, doi: 10.3348/kjr.2020.0146.

[11] G. D. Rubin *et al.*, "The role of chest imaging in patient management during the COVID-19 pandemic: A multinational consensus statement from the Fleischner Society," *Radiology*, vol. 7, Apr. 2020, Art. no. 201365, doi: 10.1148/radiol.2020201365.

[12] M. Mossa-Basha *et al.*, "Policies and guidelines for COVID-19 preparedness: Experiences from the University of Washington," *Radiology*, vol. 8, Apr. 2020, Art. no. 201326, doi: 10.1148/radiol.2020201326.

[13] L. Wang and A. Wong, "COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," 2020, *arXiv:2003.09871*.

[14] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng, "CheXNet: Radiologist-level pneumonia detection on chest X-Rays with deep learning," Nov. 2017, *arXiv:1711.05225*.

[15] J. Zhang, Y. Xie, G. Pang, Z. Liao, J. Verjans, W. Li, Z. Sun, J. He, Y. Li, C. Shen, and Y. Xia, "Covid-19 screening on chest X-ray images using deep learning based anomaly detection," 2020, *arXiv:2003.12338*.

[16] S.-H. Gao, M.-M. Cheng, K. Zhao, X.-Y. Zhang, M.-H. Yang, and P. Torr, "Res2Net: A new multi-scale backbone architecture," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 2, pp. 652–662, Feb. 2021.

[17] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans. Neural Netw.*, vol. 20, no. 1, pp. 61–80, Dec. 2009.

[18] J. Bullock, A. Luccioni, K. H. Pham, C. S. N. Lam, and M. Luengo-Oroz, "Mapping the landscape of artificial intelligence applications against COVID-19," 2020, *arXiv:2003.11336*.

[19] P. Rajpurkar *et al.*, "CheXNet: Radiologistlevel pneumonia detection on chest X-rays with deep learning," 2017, *arXiv:171105225*.

[20] H. Wang and Y. Xia, "ChestNet: A deep neural network for classification of thoracic diseases on chest radiography," 2018, *arXiv:1807.03058*.

[21] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Phys. Eng. Sci. Med.*, vol. 43, no. 2, pp. 635–640, Jun. 2020.

[22] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: A framework of deep learning classifiers to diagnose COVID-19 in X-ray images," 2020, *arXiv:2003.11055*.

[23] P. K. Sethy and S. K. Behera, "Detection of coronavirus disease (COVID-19) based on deep features," *Preprints*, 2020, Art. no. 2020030300, doi: 10.20944/preprints202003.0300.v1.

[24] A. E. Hassanien, L. N. Mahdy, K. A. Ezzat, H. H. Elmousalami, and H. A. Ella, "Automatic X-ray COVID-19 lung image classification system based on multi-level thresholding and support vector machine," *medRxiv*, 2020.

[25] O. Gozes, M. Frid-Adar, H. Greenspan, P. D. Browning, H. Zhang, W. Ji, A. Bernheim, and E. Siegel, "Rapid AI development cycle for the coronavirus (COVID-19) pandemic: Initial results for automated detection & patient monitoring using deep learning CT image analysis," 2020, *arXiv:2003.05037*.

[26] G. Nagpal, S. S. Usmani, S. K. Dhanda, H. Kaur, S. Singh, M. Sharma, and G. P. S. Raghava, "Computer-aided designing of immunosuppressive peptides based on IL-10 inducing potential," *Sci. Rep.*, vol. 7, no. 1, p. 42851, Mar. 2017.

[27] P. B. Timmons and C. M. Hewage, "HAPPENN is a novel tool for hemolytic activity prediction for therapeutic peptides which employs neural networks," *Sci. Rep.*, vol. 10, no. 1, p. 10869, Dec. 2020, doi: 10.1038/s41598-020-67701-3.

[28] C. R. Munteanu, M. Gestal, Y. G. Martínez-Acevedo, N. Pedreira, A. Pazos, and J. Dorado, "Improvement of epitope prediction using peptide sequence descriptors and machine learning," *Int. J. Mol. Sci.*, vol. 20, no. 18, p. 4362, Sep. 2019, doi: 10.3390/ijms20184362.

[29] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, M. Cai, J. Yang, Y. Li, X. Meng, and B. Xu, "A deep learning algorithm using CT images to screen for corona virus disease (COVID-19)," *medRxiv*, 2020.

[30] M. Fu, S.-L. Yi, Y. Zeng, F. Ye, Y. Li, X. Dong, Y.-D. Ren, L. Luo, J.-S. Pan, and Q. Zhang, "Deep learning-based recognizing COVID-19 and other common infectious diseases of the lung by chest ct scan images," *medRxiv*, 2020.

[31] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Y. Chen, J. Su, G. Lang, Y. Li, H. Zhao, K. Xu, L. Ruan, and W. Wu, "Deep learning system to screen coronavirus disease 2019 pneumonia," 2020, *arXiv:2002.09334*.

[32] L. Franceschi, M. Niepert, M. Pontil, and X. He, "Learning discrete structures for graph neural networks," 2019, *arXiv:1903.11960*.

[33] S. Akira, S. Uematsu, and O. Takeuchi, "Pathogen recognition and innate immunity," *Cell*, vol. 124, no. 4, pp. 783–801, Feb. 2006.

[34] T. Kawai and S. Akira, "The role of pattern-recognition receptors in innate immunity: Update on toll-like receptors," *Nature Immunol.*, vol. 11, no. 5, pp. 373–384, May 2010.

[35] D. Goubau, M. Schlee, S. Deddouche, A. J. Pruijssers, T. Zillinger, M. Goldeck, C. Schuberth, A. G. Van Der Veen, T. Fujimura, J. Rehwinkel, J. A. Iskarpatyoti, W. Barchet, J. Ludwig, T. S. Dermody, G. Hartmann, and C. Reis e Sousa, "Antiviral immunity via RIG-I-mediated recognition of RNA bearing 5'-diphosphates," *Nature*, vol. 514, no. 7522, pp. 372–375, Oct. 2014.

[36] B. K. Davis, R. A. Roberts, M. T. Huang, S. B. Willingham, B. J. Conti, W. J. Brickey, B. R. Barker, M. Kwan, D. J. Taxman, M.-A. Accavitti-Loper, J. A. Duncan, and J. P.-Y. Ting, "Cutting edge: NLRC5-dependent activation of the inflammasome," *J. Immunol.*, vol. 186, no. 3, pp. 1333–1337, Feb. 2011.

[37] H. Van Gorp, A. Kuchmiy, F. Van Hauwermeiren, and M. Lamkanfi, "NOD-like receptors interfacing the immune and reproductive systems," *FEBS J.*, vol. 281, no. 20, pp. 4568–4582, Oct. 2014.

[38] H. Hara, K. Tsuchiya, I. Kawamura, R. Fang, E. Hernandez-Cuellar, Y. Shen, J. Mizuguchi, E. Schweighoffer, V. Tybulewicz, and M. Mitsuyama, "Phosphorylation of the adaptor ASC acts as a molecular switch that controls the formation of speck-like aggregates and inflammasome activity," *Nature Immunol.*, vol. 14, no. 12, pp. 1247–1255, Dec. 2013.

[39] D. Strasser, K. Neumann, H. Bergmann, M. J. Marakalala, R. Guler, A. Rojowska, K.-P. Hopfner, F. Brombacher, H. Urlaub, G. Baier, and G. D. Brown, "Syk kinase-coupled C-type lectin receptors engage protein kinase C-δ to elicit Card9 adaptor-mediated innate immunity," *Immunity*, vol. 36, no. 1, pp. 32–42, 2012.

[40] H. Hemmi, O. Takeuchi, T. Kawai, T. Kaisho, S. Sato, H. Sanjo, M. Matsumoto, K. Hoshino, H. Wagner, K. Takeda, and S. Akira, "A toll-like receptor recognizes bacterial DNA," *Nature*, vol. 408, no. 6813, pp. 740–745, Dec. 2000.

[41] L. Sun, J. Wu, F. Du, X. Chen, and Z. J. Chen, "Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway," *Science*, vol. 339, no. 6121, pp. 786–791, Feb. 2013.

[42] S. Jensen and A. R. Thomsen, "Sensing of RNA viruses: A review of innate immune receptors involved in recognizing RNA virus invasion," *J. Virol.*, vol. 86, no. 6, pp. 2900–2910, Mar. 2012.

[43] C. Gallego-Marin, J. E. Schrum, W. A. Andrade, S. A. Shaffer, L. F. Giraldo, A. M. Lasso, E. A. Kurt-Jones, K. A. Fitzgerald, and D. T. Golenbock, "Cyclic GMP–AMP synthase is the cytosolic sensor of plasmodium falciparum genomic DNA and activates type I IFN in malaria," *J. Immunol.*, vol. 200, no. 2, pp. 768–774, Jan. 2018.

[44] D. Y. Ma and M. S. Suthar, "Mechanisms of innate immune evasion in re-emerging RNA viruses," *Current Opinion Virol.*, vol. 12, pp. 26–37, Jun. 2015.

[45] T. Nelemans and M. Kikkert, "Viral innate immune evasion and the pathogenesis of emerging RNA virus infections," *Viruses*, vol. 11, no. 10, p. 961, Oct. 2019.

[46] L. B. Ivashkiv and L. T. Donlin, "Regulation of type I interferon responses," *Nature Rev. Immunol.*, vol. 14, no. 1, pp. 36–49, Jan. 2014.

[47] L. Cao, Y. Ji, L. Zeng, Q. Liu, Z. Zhang, S. Guo, X. Guo, Y. Tong, X. Zhao, C.-M. Li, Y. Chen, and D. Guo, "P200 family protein IFI204 negatively regulates type I interferon responses by targeting IRF7 in nucleus," *PLOS Pathogens*, vol. 15, no. 10, Oct. 2019, Art. no. e1008079.

[48] S. Gordon, "Pattern recognition receptors: Doubling up for the innate immune response," *Cell*, vol. 111, no. 7, pp. 927–930, 2002.

[49] X. Lu, J. Pan, J. Tao, and D. Guo, "SARS-CoV nucleocapsid protein antagonizes IFN-β response by targeting initial step of IFN-β induction pathway, and its C-terminal region is critical for the antagonism," *Virus Genes*, vol. 42, no. 1, pp. 37–45, Feb. 2011.

[50] Y. Liu, Z. Zhang, X. Zhao, R. Yu, X. Zhang, S. Wu, J. Liu, X. Chi, X. Song, L. Fu, Y. Yu, L. Hou, and W. Chen, "Enterovirus 71 inhibits cellular type I interferon signaling by downregulating JAK1 protein expression," *Viral Immunol.*, vol. 27, no. 6, pp. 267–276, Aug. 2014.

[51] A. Okumura, P. M. Pitha, A. Yoshimura, and R. N. Harty, "Interaction between Ebola virus glycoprotein and host toll-like receptor 4 leads to induction of proinflammatory cytokines and SOCS1," *J. Virol.*, vol. 84, no. 1, pp. 27–33, Jan. 2010.

[52] E.-K. Pauli, M. Schmolke, T. Wolff, D. Viemann, J. Roth, J. G. Bode, and S. Ludwig, "Influenza a virus inhibits type I IFN signaling via NF-κB-dependent induction of SOCS-3 expression," *PLoS Pathogens*, vol. 4, no. 11, Nov. 2008, Art. no. e1000196.

[53] C. Kaewraemruaen, P. Ritprajak, and N. Hirankarn, "Dendritic cells as key players in systemic lupus erythematosus," *Asian Pacific J. Allergy Immunol.*, vol. 38, no. 4, pp. 225–232, 2020, doi: 10.12932/AP-070919-0639.

[54] W. Zapata, W. Aguilar-Jiménez, Z. Feng, A. Weinberg, A. Russo, N. Potenza, H. Estrada, and M. T. Rugeles, "Identification of innate immune antiretroviral factors during *in vivo* and *in vitro* exposure to HIV-1," *Microbes Infection*, vol. 18, no. 3, pp. 211–219, Mar. 2016.

[55] X. Zhu, Y. Wang, H. Zhang, X. Liu, T. Chen, R. Yang, Y. Shi, W. Cao, P. Li, Q. Ma, Y. Zhai, F. He, G. Zhou, and C. Cao, "Genetic variation of the human α-2-Heremans-Schmid glycoprotein (AHSG) gene associated with the risk of SARS-CoV infection," *PLoS ONE*, vol. 6, no. 8, Aug. 2011, Art. no. e23730.

[56] D. Kaur, S. Patiyal, N. Sharma, S. S. Usmani, and G. P. S. Raghava, "PRRDB 2.0: A comprehensive database of pattern-recognition receptors and their ligands," *Database*, vol. 2019, pp. 1–10, Jan. 2019.

[57] B. JD, J. LJ, N. Blom, H. G. Von, and S. Brunak, "Feature-based prediction of non-classical and leaderless protein secretion," *Protein Eng. Des. Selection*, vol. 17, pp. 349–356, Apr. 2004, doi: 10.1093/protein/gzh037.

[58] A. Garg, "A machine learning based method for the prediction of secretory proteins using amino acid composition, their order and similarity-search," *Silico Biol.*, vol. 8, pp. 40–129, Jan. 2008.

[59] M. Roberts, D. Driggs, M. Thorpe, J. Gilbey, M. Yeung, S. Ursprung, A. I. Aviles-Rivero, C. Etmann, C. McCague, L. Beer, J. R. Weir-McCall, Z. Teng, E. Gkrania-Klotsas, J. H. F. Rudd, E. Sala, and C.-B. Schönlieb, "Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans," *Nature Mach. Intell.*, vol. 3, no. 3, pp. 199–217, Mar. 2021, doi: 10.1038/s42256-021-00307-0.

[60] The UniProt Consortium, "UniProt: The universal protein knowledgebase," *Nucleic Acids Res.*, vol. 17, pp. 158–169, Jan. 2004, doi: 10.1093/nar/gkh131.

[61] *PyTorch-Lightning*. Accessed: Jan. 1, 2021. [Online]. Available: https://github.com/PyTorchLightning/pytorch-lightning

SHAGUFTA HENNA (Senior Member, IEEE) received the Ph.D. degree in computer science from the University of Leicester, U.K., in 2013. From 2018 to 2019, she was a Postdoctoral Researcher with the Telecommunication Software and Systems Group, Waterford Institute of Technology, Waterford, Ireland. She is currently a Lecturer with the Letterkenny Institute of Technology, County Donegal, Ireland. Her current research interests include big data analytics, distributed deep learning, and machine learning-driven network optimization. She is an Associate Editor of IEEE Access, *EURASIP Journal on Wireless Communications and Networking*, IEEE Future Directions, and *Human-centric Computing and Information Sciences* (Springer).